



Machine Readable Archives



BULLETIN

ISSN 0821-3658

The Appraisal of Machine Readable Data Files

The Machine Readable Archives performs archival functions for machine readable (EDP) records of long-term historical or research value generated by or for the federal government and for such information of national significance to Canada generated by the private sector. The appraisal function is one of the most critical activities in the archival process, as it is through the determination of the value of the records that they are preserved or destroyed. All records are of value to someone at some time; not all records can be preserved. It is the archivist's job to judge the value of the records relative to the information they contain, the event(s) they document and their potential use by the research community. To perform this role, the archivist must have an in-depth knowledge of the records. Decisions on the value of the records must be based on their current administrative use, their evidential and informational value, and their relationship to other records.

Appraising EDP records involves many of the same difficulties as exist for more traditional archival material as well as new ones resulting from the medium. Information, which in paper form is too cumbersome to manage, can be manipulated quickly in EDP form, thus resulting in new sources for research. The linkage possibilities with other EDP data and paper records become important considerations in the evaluation of the data. Thus, the medium permits new research possibilities that must be considered in the selection of records.

The MRA appraises EDP records in a two-step process: content analysis and technical analysis.

Content Analysis

Content analysis is the evaluation of the information contained in the records and involves the more traditional appraisal techniques: the determination of the archival value of the records based on their research value. The research or historical value of EDP records is determined by examining the evidential, informational, and legal value of the records. EDP records that have evidential value are those that provide "evidence" or testimony of the existence of a department, its organization, and its functions or activities. In order to judge the evidential value, the archivist must have an in-depth knowledge of the internal organization of the department or agency, the programs it supports, and the activities it carries out. The

MRA has acquired training program data and immigration data from the Canada Employment and Immigration Commission. Such data were created by the department in support of the programs and services administered under its responsibilities. These data therefore relate directly to the mandate of the department and were appraised as having evidential value.

The determination of the informational value of EDP records is, of course, based on the information contained in the records. EDP data are collections of information relating to persons, things, or activities. General points to consider in the determination of informational value are the uniqueness of the information, the form, and the importance. The uniqueness of the data relates to whether or not the information can be found elsewhere and in what form. If the records are in machine readable form only, do duplicate copies exist? Are these data part of a large file? Are the data on the micro-level or are they aggregated? Non-aggregated data are of more research value than aggregate data; a master file has more research potential than an extract from the master file. Such considerations must be investigated during the appraisal of the records.

The determination of the informational value of the records is the most subjective aspect in the appraisal process. It is at this point that the archivist's knowledge of current research trends and methods and the contact with the research community play a vital part in the decision to keep or destroy the data. Knowledge of the use(s) to which information was put (primary use) and the use(s) to which information can be put (secondary use) is essential. The special characteristics of EDP records such as their manipulability and the linkage possibilities increase their informational value.

Technical Analysis

EDP records cannot be appraised solely on their content; they must be examined in terms of their technical characteristics as well. An obvious example of the importance of the technical aspect is the medium itself. If the data reside on magnetic tape and the tape is damaged to the point where data cannot be read, then appraisal of the information is impossible. The internal arrangement of the data may be so complex that servicing the records would be too costly or impossible. The size and periodicity of the data are also

important considerations as is the data's software dependency. Many of the technical considerations relate in the long run to the cost of acquisition, processing, conservation, and servicing of the records. In theory, the costs associated with technical problems should not play a role in the appraisal process; in reality, they cannot be neglected as the division only has a certain amount of funds. The archivist must be aware of how the EDP data came into existence, the difficulties involved in the processing and conservation of the records, and the knowledge required to service the records. A knowledge of data processing techniques and their evolution is required in the technical analysis.

Summary

The appraisal process is complex and requires that the divisional archivists have an in-depth knowledge of the contents of the data, the reasons for their creation, how they were created, their present use, and ways in which they can be used in the future. At the same time, the archivist must be aware of existing EDP technology and future developments. Such responsibilities require that archivists consult with people knowledgeable in these fields. As was said by Schellenberg, a leading American archivist,

"... Appraisals of records should not be based on intuition or arbitrary supposition of value; they should be based instead on thorough analyses of the documentation bearing on the matter to which the records pertain. Analysis is the essence of archival approval."^{*}

* Schellenberg, Theodore R. *The Appraisal of Modern Records*. Washington: National Archives, 1956, p. 65.

"Towards a Canadian Electronic Heritage"

The MRA of the Public Archives of Canada and the Social Policy and New Services Directorate of the Department of Communications have undertaken a joint project. The purpose of the project is twofold: to develop a strategy for coordinating policies in the communications/cultural sector respecting the management, deposit, and maintenance of publically funded machine readable data files of archival and research value; and to develop a strategy for creating a nationally accessible, up-to-date, public, on-line inventory of Canadian data files and data bases.



"Canadian Electronic Heritage"

cont'd from page 1

The first phase of the project, which was completed at the end of April, involved consultations or discussions with key individuals across the country, in both the public and private sectors, known to have an interest or expertise in these areas.

The second phase of the project involved a one-day session in Vancouver on 6 June during the Learned Societies' Conference. The purpose of the session was to discuss the various options that were suggested during the consultative process. The basis of the discussion was the consultant's report entitled "Towards a Canadian Electronic Heritage." The day was divided into five main topics with an introduction at the beginning of the day by the sponsors of the session, Mary Meloshe and Harold Naugler and the consultant, Jake Knoppers. The topics addressed by the participants were: Data Archives/Libraries: Available Infrastructures and Possible Roles; Infrastructure Coordination and Cooperation at the National Level; Agency Perspectives; Scholarship and Access: Views of Researchers and Users; Ensuring an Electronic Cultural Heritage: Agency Perspectives; and *Quo Vadis*. Although no specific conclusions were reached, the session brought together for the first time in a number of years, those involved in the creation, collection, distribution, and funding of machine readable data and allowed all the participants to discuss the existing problems and possible solutions. Over the next few months, officers in the Department of Communications and the Public Archives will be reviewing the issues raised in the discussion paper, considering the suggestions and recommendations made throughout the consultative process and during the Vancouver forum, and discussing the appropriate policy responses at the federal level.

Anyone who would like more information on the session or a copy of the report should contact the Documentation and Public Service Section of the Machine Readable Archives.

Search for Data Files on the Role of Women in Canadian Society

On 9 and 10 December 1982, a Conference on Data Requirements to Support Research into Women and the Canadian Economy was held at the University of Manitoba. The Conference was organized by the staff of the Institute for Social and Economic Research at the University, with funding provided by the Social Sciences and Humanities Research Council of Canada. Many of the leading Canadian researchers in this field of study presented papers. A set of the Conference proceedings can be obtained by writing to Maureen Woodrow, Social Sciences and Humanities Research Council of Canada, 255 Albert Street, Ottawa, Ontario, K1P 6G4.

Both in their papers and during the various discussions, the conference participants referred repeatedly to the lack of machine readable data that are available in this strategic research area. There was unanimous consensus that more research needed to be undertaken in order to identify existing sources of data. There was also a need for additional questions concerning women in some of the surveys currently undertaken by the national statistical agency, particularly the Labour Force Surveys. And finally, new machine readable data files must be funded and used to build up a useful data base in the field.

Over the past two years, the Machine Readable Archives has developed an automated inventory containing information on computer systems and individual data files that exist in federal government departments and agencies and in the private sector (including data archives and libraries, research institutions, as well as data in private hands). Most of the public sector information is the result of surveys that have been undertaken in the federal government. Although it is far from comprehensive at the present time, it does serve as a useful starting point to determine what exists in machine readable form in the federal government. Most of the private sector information in the data base has come from the printed catalogues of the former Data Clearing House for the Social Sciences. Some of the data files described in the automated inventory have been acquired by the MRA, but in most cases the data are still located in the creating institutions. Divisional staff will undertake a search of the data base to locate data on a particular subject. This service is provided free of charge.

At the present time, there are few sources of information in the data base on the role of women in Canadian society, a situation that we would like to correct. Therefore, we would very much appreciate receiving information on any data files in this field of study that individuals or institutions might have in their possession. Please contact the Documentation and Public Service Section, Machine Readable Archives, Public Archives of Canada, 395 Wellington Street, Ottawa, Ontario, K1A 0N3, or phone (613) 593-7772.

IASSIST '83

The 1983 IASSIST (International Association for Social Science Information Service and Technology) Conference was held in Philadelphia from 19 May to 22 May. Despite cutbacks and lack of travel funds, approximately eighty people participated in the Conference. Workshops were held the day prior to the Conference. This year there were three workshops: microcomputers; complex datasets; and planning a data library. The Conference got underway with an excellent presentation by Professor Robert Kraft from the Department of Religious Studies at the University of Pennsylvania. His address on the use of computers in the comparison of religious texts was very informative and well received by a group whose main interests to date have been numeric data. A Blue Ribbon Panel on the "Development of an Information Network for Locating and Describing Machine Readable Data Files" brought together representatives from the United States, Canada and Europe and provided an interesting, if not somewhat controversial, discussion on this subject. There were three main tracks for the sessions: the Hardware and Software track dealt specifically with the issues surrounding microcomputers and data storage; the track on Data Files focused on census related files and complex data files; and the Data Services track dealt with indexing for access to social science data and planning and resource management systems.

All of the members of the Conference committees did a marvellous job in providing interesting sessions and the usual IASSIST hospitality. Next year's Conference will be held in Ottawa in mid-May. Anyone interested in the Conference should contact the MRA.

New Acquisitions

The division has acquired a series of data files from the Department of Industry, Trade and Commerce on *Vacation and Travel by Canadians*. The files contain data from surveys conducted in 1967, 1968, 1971, 1974 to 1979, and 1981. The purpose of the surveys was to obtain descriptions of holidays and vacation trips taken by Canadians in Canada, the United States, and elsewhere. The files will be processed during the year.

d'obtenir des informations sur les vacances et les voyages des Canadiens au Canada, aux États-Unis et dans d'autres pays. Ces fichiers seront traités durant l'année.

Recherche de fichiers informatisés sur le rôle des femmes dans la société canadienne

Les 9 et 10 décembre derniers, il y a eu à l'université du Manitoba une conférence sur

les données requises pour effectuer des recherches sur les femmes et l'économie canadienne. Cette conférence était organisée par le personnel de l'Institut for Social and Economic Research de l'université et financée par le Conseil de recherches en sciences humaines du Canada. De nombreux chercheurs canadiens réputés dans ce domaine y ont présenté des exposés. On peut obtenir les comptes rendus de la conférence en écrivant à Maureen Woodrow, Conseil de recherches en sciences humaines, 255, rue Albert, Ottawa (Ontario) K1P 6G4.

Durant les exposés et les discussions, il a été question à maintes reprises de l'absence de données automatisées dans ce domaine de recherche stratégique. Tous les participants s'entendaient sur la nécessité d'intensifier les recherches pour trouver des sources de données existantes. Ils ont aussi recommandé d'ajouter des questions sur les femmes dans certaines enquêtes actuellement menées par l'organisme national de statistiques, notamment dans les enquêtes sur la main-d'œuvre. Et enfin, il faudrait selon eux financer de nouveaux fichiers automatisés et les utiliser pour constituer une base de données utile dans le domaine.

La division a acquis une série de fichiers informatisés du ministère de l'Industrie et du Commerce sur les vacances et les déplacements des Canadiens. Ils renferment des données des enquêtes menées en 1967, 1968, 1971, de 1974 à 1979, et en 1981, dans le but de la division interroger la base de données blissements qui les ont créés. Le personnel ordi- niques mais, dans la plupart des cas, les données se trouvent encore dans les éta- ventaire ont été acquis par les Archives ciales. Certains des fichiers décrits dans l'in- coordination des données en sciences so- catalogues imprimés de l'ancien centre de ments sur le secteur privé sont tirés des vernement fédéral. La plupart des renseigne- pour connaître les fonds automatisés du gou- tant, elle constitue un bon point de départ qu'elle soit loin d'être exhaustive pour l'in- le secteur public provient d'enquêtes effec- tées dans l'administration fédérale. Bien- liers). La majeure partie de l'information sur la recherche et ceux détenus par des particu- et des bibliothèques, des établissements de m- nistres et organismes fédéraux et du secteur les fichiers automatisés individuels des mi- taire automatisé contenant de l'information au cours des deux dernières années un inven- Les Archives ordi- niques ont dressé

Nouvelles acquisitions

Pour obtenir d'autres renseignements sur la séance ou un exemplaire du rapport, veuillez vous adresser à la Section de la docu- mentation et du service au public des Archi- ves ordi- niques.

La deuxième étape du projet consistait à tenir une séance d'une journée, le 6 juin à Vancouver dans le cadre de la conférence des Sociétés savantes, pour discuter des diverses options suggérées durant le processus de consultation. La discussion était basée sur le rapport de l'expert-conseil intitulé «Towards a Canadian Electronic Heritage». La séance a été ouverte par les organisateurs Mary Melosh et Harold Nangle, et l'expert-con- seil Jake Knoppers. Elle portait sur cinq sujets principaux : dépôts d'archives et bibliothèques ; infrastructures existantes et rôles pos- sibles ; coopération et coordination de l'intra- structure au niveau national ; point de vue des agences gouvernementales ; érudition et accès ; opinions des chercheurs et des utilis- teurs ; préparation d'un patrimoine culturel informatisé ; point de vue des agences gou- vernementales ; et *Quo Vadis*. Les partici- pants n'en sont pas arrivés à des conclusions précises, mais cette séance a tout de même réuni, pour la première fois depuis plusieurs années, des personnes travaillant à la créa- tion, à la collecte, à la distribution et au finan- cement de données ordi- niques et leur a permis de discuter des problèmes existants et des solutions possibles. Dans les prochains mois, des représentants du ministère des Communications et des Archives publiques passeront en revue les questions soulevées dans le document de travail, examineront les suggestions et les recommandations formu- lées au cours du processus de consultation et du colloque de Vancouver, et discuteront en- suite des politiques à adopter au niveau fédéral.

Conférence de 1983 de l'IASIST

La conférence de 1983 de l'IASIST (As- sociation internationale pour les services et techniques d'information en sciences so- ciales) a eu lieu à Philadelphie du 19 au 22 mai. Malgré la réduction des dépenses de déplacement, environ quatre-vingts person- nes y ont participé. Cette conférence avait été précédée la veille de trois ateliers : les micro- ordinateurs, les ensembles de données com- plexes et la planification d'une bibliothèque gérant des données. La conférence a débuté par un excellent exposé du professeur Robert Kraft, du département des études religieuses à l'université de Pennsylvanie. Ses propos sur l'utilisation d'ordinateurs pour comparer des textes religieux étaient très enrichissants et ont été bien accueillis par un auditoire qui, jusqu'à maintenant, s'était surtout intéressé aux données numériques. Une discussion spéciale sur l'établissement d'un réseau d'in- formation destinée à localiser et à décrire des fichiers ordi- niques, réunissant des repré- sentants des États-Unis, du Canada et de l'Europe, a donné lieu à des échanges inté- ressants et même suscité une légère contro- verse. Les séances se divisaient en trois sections principales : la section «matériel et logiciel» portait surtout sur les questions rela- tives aux micro-ordinateurs et au stockage des données, la section «fichiers informati- sés» était axée sur les fichiers provenant des recensements et les fichiers complexes, et la section «services informatiques» traitait du répertoriage des données en sciences sociales pour les rendre plus accessibles, ainsi que de la planification et des systèmes de gestion des ressources.

Nous tenons à remercier l'IASIST de sa cordiale hospitalité et à souligner le merveilleux travail accompli par tous les membres des comités de la conférence qui ont présenté des séances très intéressantes. La conférence de l'an prochain aura lieu à Ottawa à la mi- mai. Ceux qui aimeraient y participer doivent communiquer avec les Archives ordi- niques.

Archives ordinolinges



3 1761 11550425 0

Evaluation des fichiers ordinolingués

Les Archives ordinaires remplissent

tion, et de ses fonctions ou activités. Pour pouvoir en juger, l'archiviste doit connaître à fond l'organisation interne du ministère ou de l'organisme, ses activités et les programmes qu'il appuie. Les Archives ordinaires ont acquis d'Emploi et Immigration Canada des données sur les programmes de formation et sur l'immigration. Ces données ont été créées par le ministère dans le cadre des programmes et des services dont il est responsable. Elles se rapportent donc directement au mandat du ministère et sont considérées comme ayant une valeur de témoin.

gnage.
La

La valeur informative des documents automatisés est évidemment fonction de l'information qu'ils contiennent. Les données

impossible d'évaluer l'information. La disposition interne des données peut être tellement complexe qu'il serait impossible ou trop coûteux d'entretenir les documents. La quantité et la périodicité des données sont également des facteurs importants, de même que la fiabilité du logiciel employé. Nombre des aspects techniques à considérer ont trait à long terme au coût de l'acquisition des documents, de leur traitement, de leur conservation et de leur entretien. En principe, les coûts associés aux problèmes techniques ne devraient pas être pris en considération, mais ils ne peuvent être ignorés dans les faits car les fonds budgétaires de la division sont limités. L'archiviste doit savoir comment les données informatisées ont été engendrées et comment elles doivent être entretenues, et connaître les difficultés soulevées par leur traitement et leur conservation. L'analyse technique exige une connaissance des méthodes de traitement des données et de leur évolution.

Résumé

Le processus d'évaluation est complexe et exige que les archivistes de la division continuent de fonder le contenu des données, les motifs de leur création, la façon dont elles ont été créées, leur utilisation actuelle et future. Ils doivent aussi se tenir au courant des techniques automatisées existantes et des développements à venir. Ces responsabilités les obligent à consulter des personnes bien formées dans ces domaines. Comme le disait Schellenberg, réputé archiviste américain : « . . . Les évaluations de documents ne doivent pas être basées sur l'intuition ni sur des suppositions arbitraires, mais plutôt sur des analyses approfondies de la documentation sur la question à laquelle les documents se rapportent. L'analyse est l'essence même de l'évaluation des archives. »*

* Schellenberg, Theodore R. *The Appraisal of Modern Records*. Washington, National Archives, 1956, p. 65.

«Vers un patrimoine canadien informatisé»

Les Archives ordinolingués, des Archi-
ves publiques du Canada, et la Direction des
politiques sociales et nouveaux services du
ministère des Communications ont entrepris
ensemble un projet ayant un double but : éla-
borer une stratégie pour coordonner les poli-
tiques, dans le secteur de la culture et des

liaison) accroissent leur valeur informative.

Analyse technique

Les documents automatisés ne peuvent être évalués d'après leur seul contenu, il faut aussi tenir compte de leurs caractéristiques techniques. On n'a qu'à penser par exemple au support matériel. Si les données sont sur un ruban magnétique et qu'il est endommagé au point où elles sont illisibles, il est alors

tenu et l'analyse technique.

Analyse du contenu

L'analyse du contenu est l'évaluation de l'information consignée dans les documents et fait appel aux techniques d'évaluation plus traditionnelles ; elle consiste à déterminer la valeur archivistique des documents en fonction de leur valeur documentaire. On détermine la valeur documentaire ou historique des documents automatisés en étudiant leur valeur de témoignage, informative et juridique. Les documents automatisés ont une valeur de témoignage lorsqu'ils attestent l'existence d'un ministère, de son organisation